



DePaul Law Review

Volume 66
Issue 1 *Fall 2016: Twenty-Sixth Annual DePaul
Law Review Symposium*

Article 6

Computer Source Code: A Source of the Growing Controversy Over the Reliability of Automated Forensic Techniques

Edward J. Imwinkelried

Follow this and additional works at: <https://via.library.depaul.edu/law-review>



Part of the [Law Commons](#)

Recommended Citation

Edward J. Imwinkelried, *Computer Source Code: A Source of the Growing Controversy Over the Reliability of Automated Forensic Techniques*, 66 DePaul L. Rev. (2017)

Available at: <https://via.library.depaul.edu/law-review/vol66/iss1/6>

This Article is brought to you for free and open access by the College of Law at Via Sapientiae. It has been accepted for inclusion in DePaul Law Review by an authorized editor of Via Sapientiae. For more information, please contact digitalservices@depaul.edu.

COMPUTER SOURCE CODE: A SOURCE OF THE GROWING CONTROVERSY OVER THE RELIABILITY OF AUTOMATED FORENSIC TECHNIQUES

*Edward J. Imwinkelried**

I. INTRODUCTION

We [are] . . . increasingly . . . governed by source code If we do not recognise that our . . . processes and procedures . . . are gradually becoming source code, then we risk a technological tyranny.¹

—Professor Mark Harman, Department of Computer Science, University College London²

Courts throughout the world are increasingly relying on expert testimony, notably scientific evidence, to resolve legal disputes. One Rand Corporation study of over 500 American trials found that experts presented testimony at eighty-six percent of trials.³ One commentator has suggested that in the United States, trial by jury is evolving into trial by expert.⁴

In the early 20th century, the normal pattern was that an individual expert would personally conduct by hand a single test such as a drug identification procedure. Early forensic tests were labor-intensive and sometimes slow. Today, a forensic scientist witness often does not appear at trial to describe a manual scientific analysis. Rather, the witness will testify about the results of an automated forensic technique that he or she oversaw. In a drunk driving case, the witness is rarely a

* Edward L. Barrett, Jr. Professor of Law Emeritus, University of California, Davis School of Law; former chair, Evidence Section, American Association of Law Schools; coauthor, PAUL C. GIANNELLI, EDWARD IMWINKELRIED, ANDREA ROTH & JANE CAMPBELL MORIARTY, *SCIENTIFIC EVIDENCE* (5th ed. 2012). A version of this Article was presented at the Second International Symposium on Sino Swiss Evidence Science at the University of Lausanne, Switzerland on September 6–9, 2016.

1. Mark Harman, *Why Source Code Analysis and Manipulation Will Always Be Important*, TENTH IEEE INT'L WORKING CONFERENCE ON SOURCE CODE ANALYSIS & MANIPULATION, Sept. 12–13, 2010, at 7, <http://ieeexplore.ieee.org/amp/stamp.jsp?arnumber=5601835>.

2. Mark Harman, Professor of Software Engineering, University College London, <http://www.0.cs.ucl.ac.uk/staff/mharman/>.

3. Samuel R. Gross, *Expert Evidence*, 1991 WIS. L. REV. 1113, 1119 (1991) (discussing results derived from civil trials in California in 1985 and 1986).

4. William T. Pizzi, *Expert Testimony in the US*, 145 NEW L.J. 82, 82 (1995) (describing how experts are “readily availab[le]” to testify in favor of each party).

toxicologist testifying about a manual oxidative analysis of the alcohol concentration in a blood sample that he or she personally drew from the suspect.⁵ It is far more likely that the witness will be a police officer trained to use an Intoxilyzer that samples the suspect's breath and contains a computer program that reads out an estimate of the suspect's breath alcohol concentration.⁶ In a rape case, the witness may not be a DNA analyst who personally analyzed the vaginal swab from the victim of a gang rape and opines as to the DNA profile of all the contributors to the mixed sample. Eighty-five percent of DNA laboratories in the United States use automated techniques, such as Applied Biosystems instrumentation,⁷ to conduct the initial analysis to identify the DNA profiles in a sample.⁸ Assume the results of this initial stage of the analysis indicate that there were multiple contributors with varying profiles to the genetic sample. On that assumption, the witness may testify to the result when he or she used Cybergenetics' TrueAllele computer program, which applies probabilistic genotyping⁹ to generate a statistic as to whether a particular defendant was one of the contributors.¹⁰ To enhance uniformity of analysis and partially to increase the volume of analyses that a laboratory can conduct,¹¹ there is a growing trend to automate—that is, computerized forensic analysis.

The “heart” of any computer program is its source code.¹² The source code contains all the instructions that the program needs to execute its tasks. Some have described DNA as the biological source code for life.¹³ By way of analogy, one commentator has said that the

5. See 2 PAUL C. GIANNELLI ET AL., SCIENTIFIC EVIDENCE § 20.04[b] (5th ed. 2012) (“[A] 1999 survey conducted by the College of American Pathologists revealed that [the oxidative] method is rarely used today.”).

6. See *id.* § 22.03[b][1] (describing the Intoxilyzer testing process).

7. *Id.* § 18.03[d].

8. William C. Thompson & Dan E. Krane, *DNA in the Courtroom*, in PSYCHOLOGICAL & SCIENTIFIC EVIDENCE IN CRIMINAL TRIALS § 11:23 (2003).

9. See 2 GIANNELLI ET AL., *supra* note 5, at 83–87, for a general description of the computer-based, probabilistic genotyping approach to analyzing mixed DNA samples.

10. TrueAllele is not the only program relying on probabilistic genotyping. STRMix and Forensic Statistical Tool (FST) are also such programs. See *People v. Muhammad*, No. 14-65263-FC, at *2 (Mich. Cir. Ct. Dec. 15, 2015), http://media.mlive.com/chronicle/news_impact/other/muhammad-opinion.pdf (discussing STRMix).

11. Jim Dawson, *Fighting Crime with Science*, NAT'L INST. JUST. J., Dec. 2015, at 10 (noting that forensic laboratories are dealing with “a never-ending flow of evidence,” and that “the demands on crime laboratories to process evidence faster while lowering costs have increased dramatically over the past decade”).

12. Christopher M. Mislaw, *Protecting Source Code from Disclosure During Pretrial Discovery*, 12 UTAH B.J. 39, 47 (1984).

13. Harman, *supra* note 1, at 10.

source code is the “lifeblood” of a computer program.¹⁴ The accuracy of the source code is the most fundamental guarantee of the program’s proper operation.¹⁵ The code dictates which tasks a computer program performs, how the program performs the tasks, and the order in which the program performs the tasks.¹⁶

At first blush, given the fundamental importance of a computer program’s software, one would think that when a question arose as to whether the software was operating correctly, the courts would demand a showing of the validity of the code or at least allow a litigant challenging the program to examine the code itself. Courts in the United States frequently allow civil litigants to perform such an examination. On the civil side, it is something of an aberration for a court to refuse to permit a civil litigant challenging the performance of a computer program to have an independent expert study the code.¹⁷ In many civil disputes, such as the highly publicized controversy over the Volkswagen computer program allegedly designed to defeat emissions testing,¹⁸ the content of the computer’s code is directly relevant. In the *Volkswagen* case, the central question was whether the company wrote code that was intended to mislead emission tests.

In the criminal setting, the content of computer code ordinarily has merely circumstantial relevance; the code is relevant only to the reliability of a scientific instrument that in turn has been used to generate evidence of a disputed fact against the accused. As a general proposition, in the United States, criminal discovery is more limited than civil discovery.¹⁹ More specifically, criminal courts have taken a very dif-

14. Jorge R. Roig, *Decoding First Amendment Coverage of Computer Source Code in the Age of Youtube, Facebook, and the Arab Spring*, 68 N.Y.U. ANN. SURV. AM. L. 319, 395 (2012); see DOUGLAS R. HOFSTADTER, GODEL, ESCHER, BACH: AN ETERNAL GOLDEN BRAID 290–91 (1979).

15. Ryan Christopher Fox, Comment, *Old Law and New Technology: The Problem of Computer Code and the First Amendment*, 49 UCLA L. REV. 871, 873 (2002).

16. See generally Harman, *supra* note 1.

17. Mislow, *supra* note 12, at 47 (comparing several cases in which the court allowed the litigant to access and examine the code with *Hubco Data Products Corp. v. Mgmt. Assistance, Inc.*, No. 81-1295, 1983 WL 1130, at *3 (D. Idaho Feb. 3, 1983) (Memorandum Decision denying motion to compel production of documents)).

18. Guilbert Gates et al., *Explaining Volkswagen’s Emissions Scandal*, N.Y. TIMES (July 19, 2016), http://www.nytimes.com/interactive/2015/business/international/vw-diesel-emissions-scandalexplained.html?_r=0; Russell Hotten, *Volkswagen: The Scandal Explained*, BBC (Dec. 10, 2015), <http://www.bbc.com/news/business-34324772>. If a corporation hires a computer company to design software to control some of its business or industrial operations and the corporation believes that the software does not function properly, the corporation may sue the company for breach of warranty. The corporation’s complaint will contain express allegations about the software, making the contents of the program directly relevant.

19. 1 GIANNELLI ET AL., SCIENTIFIC EVIDENCE § 3.01 (5th ed. 2012). On its face, Civil Rule 26 purports to authorize much broader discovery than Criminal Rule 16. Civil Rule 26 allows for

ferent approach to source code discovery. There have been two primary waves of criminal cases raising the issue.²⁰ The first wave crested in the early 2000s. During that period, the issue was access to the source code of the software programs for computerized infrared breath testing devices. With few exceptions,²¹ the clear majority of courts rejected defendants' requests that a defense expert be granted access to the program's source code.²² The second wave is occurring now. Today, the issue presented is gaining access to the source code for a probabilistic genotyping computer program, such as TrueAllele, that generates statistics for the DNA profiles of potential contributors to mixed DNA samples—perhaps the most daunting challenge for a DNA laboratory.²³ To date, TrueAllele has been used in approxi-

discovery of "any nonprivileged matter that is relevant to any party's claim or defense," while Criminal Rule 16 only allows discovery of certain types of evidence. *Compare* FED. R. CIV. P. 26, *with* FED. R. CRIM. P. 16.

20. There have been isolated source code cases that have not involved either breath testing devices or the TrueAllele DNA program. One such case was *United States v. Chiaradio*, 684 F.3d 265, 276–77 (1st Cir. 2012). Chiaradio was charged with possession and distribution of child pornography. *Id.* at 265. A government expert used an enhanced peer-to-peer file-sharing program (EP2P) to download files and facilitate the identification of files containing child pornography. The defense claimed that the EP2P technology had not been sufficiently tested. *Id.* at 276–77. The court held that the government's foundation was adequate even though the program's source code had not been peer reviewed. *Id.* Although there are isolated decisions such as *Chiaradio*, the vast majority of the published criminal cases relating to source code deal with either breath testing devices or probabilistic genotyping programs such as TrueAllele. *See also* *United States v. Matish*, No. 16-cv-16, 2016 U.S. Dist. LEXIS 82279 (E.D. Va. June 23, 2016) (challenging the source code for the full network investigative technique employed in child pornography cases to examine a suspect's computer).

21. *See* *State v. Bjorkland*, 924 So. 2d 971, 973 (Fla. Dist. Ct. App. 2006); *State v. Underdahl*, 767 N.W.2d 677, 683–84 (Minn. 2009); *In re Comm'r of Public Safety*, 735 N.W.2d 706, 712–13 (Minn. 2007); *State v. Chun*, 923 A.2d 226, 226–27 (N.J. 2007), *cert. denied*, 555 U.S. 825 (2008).

22. *See, e.g.,* *Moe v. State*, 944 So. 2d 1096, 1097 (Fla. Dist. Ct. App. 2006); *State v. Underdahl*, 749 N.W.2d 117, 120–21 (Minn. Ct. App. 2008), *aff'd in part, rev'd in part*, 767 N.W.2d 677 (Minn. 2009) (affirming the district court's denial of production of computer code); *State v. Burnell*, No. MV06479034S, 2007 WL 241230, at *2 (Conn. Super. Ct. Jan. 18, 2007); *State v. Walters*, No. DBDMV050340997S, 2006 WL 785393, at *1 (Conn. Super. Ct. Feb. 15, 2006); *People v. Cialino*, 831 N.Y.S.2d 680, 681–82 (Crim. Ct. 2007).

23. Katherine L. Moss, Note, *The Admissibility of TrueAllele: A Computerized DNA Interpretation System*, 72 WASH. & LEE L. REV. 1033, 1060–62 (2015) ("The TrueAllele computer program relies on a form of statistical analysis called probabilistic genotyping. Probabilistic genotyping involves applying the information derived from DNA profiles to complex mathematical formulas known as algorithms. The algorithms compare different statistical models to the actual data and weigh the probability that the model matches the data Specifically, TrueAllele relies on a class of algorithms derived from a Bayesian statistical analysis called Monte Carlo-Markov Chain (MCMC) modeling. The MCMC statistical approach has been used in a variety of situations to successfully model many complex data sets."). The developer of TrueAllele, Dr. Mark Perlin, claims that TrueAllele yields more reliable statistics in DNA mixture cases than previously employed approaches. *See* Mark Perlin, *Inclusion Probability for DNA Mixtures Is a Subjective One-Sided Match Statistic Unrelated to Identification Information*, J. PATHOLOGY INFORMATICS (Oct. 28, 2015), <http://www.jpathinformatics.org/article.asp?issn=2153-3539;year=>;

mately 500 cases.²⁴ As in cases involving infrared breath testing devices, defendants throughout the country have sought to discover TrueAllele's source code in order to have the code evaluated by defense experts.²⁵ The courts' responses to requests for the source code of TrueAllele has been even more uniformly negative than the previous requests for access to the source codes of infrared breath testing instruments.²⁶ In fact, "[c]ourts in at least seven states" have rebuffed defendants' requests.²⁷

This Article argues that although the prosecution may lay the foundation for admitting testimony based on probabilistic genotyping programs, such as TrueAllele, without presenting expert testimony about the program's source code, in some circumstances a criminal accused should have the right to access to the program's source code to assess its validity. Part II of this Article explains the concept of computer source code²⁸ and reviews the multi-step process that leads to the creation of such code. Part III of the Article describes the prior case law admitting testimony based on such programs while denying the de-

volume=;issue=;page=;epage=;aulast=Perlin (criticizing the use of the combined probability of inclusion (CPI), employed "by crime labs for over 15 years"); *see also* Seth Augenstein, *DNA Mixture Calculation Method Just 'Random Number Generator,' Says New Study*, FORENSIC MAG. (Nov. 16, 2015, 12:50 PM), <http://www.forensicmag.com/article/2015/11/dna-mixture-calculation--random-number-generator-says-new-study> (describing the unreliable nature of DNA testing in laboratories). In PRES. COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT – FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (2016), the President's Council of Advisors on Science and Technology is sharply critical of the subjective methods used in the past to analyze complex DNA mixtures. The Council asserts that the approach used in the past "has been an inadequately specified—and thus inappropriately subjective—method. As such, the method is clearly not foundationally valid." *See id.* at 78. There are at least eight probabilistic genotyping programs, including TrueAllele, STRmix, and FST. *Id.*

24. Augenstein, *DNA Mixture*, *supra* note 23.

25. *E.g.*, Paula Reed Ward, *Judge to Allow DNA Evidence in Deaths of Wolfe Sisters*, PITT. POST-GAZETTE (Oct. 30, 2015, 12:00 AM), <http://www.post-gazette.com/local/city/2015/10/30/Judge-to-allow-DNA-evidence-in-deaths-of-East-Liberty-Wolfe-sisters/stories/201510300182> ("[A]ttorneys have asked Common Pleas Judge Jill E. Rangos to require TrueAllele's creator, Dr. Mark Perlin, to release his source code so they can independently verify his results.").

26. *See, e.g.*, *Commonwealth v. Robinson*, No. CC 201307777 (Pa. Ct. C.P. Allegheny Cty. Feb. 4, 2016) (discussing the admissibility of test results without disclosing source code in previous trials).

27. Joe Palazzolo, *Judge Denies Access to Source Code for DNA Software Used in Criminal Cases*, WALL ST. J. (Feb. 5, 2016, 10:56 AM), <http://blogs.wsj.com/law/2016/02/05/judge-denies-access-to-source-code-for-dna-software-used-in-criminal-cases/> (discussing the recent ruling of Judge Jill Rangos in a prosecution of Michael Robinson in Pennsylvania); Jonathan D. Silver, *Stay, Immediate Appeal Sought for Attorneys Seeking Access to DNA Analysis Software*, PITT. POST-GAZETTE (Jan. 5, 2016, 10:25 PM), <http://www.post-gazette.com/city/01/05/-DNA-analysis-software-in-potential-capital-case-ask-for-stay-immediate-appeal/stories/201601050217>.

28. *See infra* notes 34–73 and accompanying text.

fense access to the program's source code.²⁹ In particular, this Part reviews the infrared breath testing devices cases and decisions on TrueAllele.

While Parts II and III of the Article are descriptive, Part IV presents a critical evaluation of the issues. Initially, this Part addresses the question of whether the prosecution should be permitted to introduce expert testimony based on a computerized technique without presenting foundational testimony about the validity of the program's source code.³⁰ This Part concludes that the prosecution may do so. Part IV then turns to the issue of whether there are circumstances in which the defense ought to have access to the source code.³¹ This Part considers both the defense's alleged need for access and the manufacturer's countervailing interest in shielding a source code, which is often a valuable trade secret. Part IV argues that when the available validation studies for the program omit a condition that (1) is present in the pending case and (2) could materially affect the performance of the program, the defense should have a limited right of access.³² Part IV concludes by proposing a specific procedure that trial judges can employ in deciding whether and how to grant defendants access.³³

II. AN EXPLANATION OF COMPUTER SOURCE CODE

Computer code is an enormous, complex topic.³⁴ This Part of the Article presents a brief, simplified overview of the process. The overview explains why the focal point of the controversy over TrueAllele is the program's source code rather than any other aspect of the program.

A. *The General Concept of Source Code*

Source code dictates which tasks a program performs, how the program performs the tasks, and the sequence in which the program performs the tasks. A computer accomplishes tasks³⁵ by executing³⁶ the

29. See *infra* notes 74–107 and accompanying text.

30. See *infra* notes 108–54 and accompanying text.

31. See *infra* notes 155–99 and accompanying text.

32. See *infra* notes 200–28 and accompanying text.

33. See *infra* notes 229–35 and accompanying text.

34. See JAMES GRIMMELMANN, *INTERNET LAW: CASES & PROBLEMS* 25 (6th ed. 2016) (“There are thousands of programing languages.”).

35. See Trevor J. Foster & Seth A. Northrop, *A Lawyer's Guide to Source Code Discovery*, FED. L., Feb. 2011, at 42.

36. Harman, *supra* note 1, at 8–9.

orders contained in the code;³⁷ simply stated, the program operates by carrying out those orders.³⁸ The orders take the form of algorithms embedded in the code.³⁹

B. The Multi-Step Process of Creating Source Code

The process begins when an employee or team of employees develops the concept for marketing a product to perform a particular task. The starting point is producing a document specifying the marketing requirements—the functions that the company will eventually advertise the product as performing.⁴⁰

The marketing team passes this document on to the programming group members, who attempt to identify the technical requirements for the product.⁴¹ The group formulates a general design including a functional flow that will enable the user to employ the product to perform all tasks identified by the marketing team.⁴²

Next, the group converts the general design concept into a more detailed specification document.⁴³ In this phase, the group begins developing a deeper, more comprehensive understanding of everything that the program must contain in order to perform the tasks identified by the marketing team.⁴⁴

Now the work of the codewriters⁴⁵ starts in earnest. Initially, they must choose a language in which to write the code. Over the decades codewriters have developed and used a large number of codes, including BASIC, C, C++, COBOL, Fortran, Java, HTML, LISP, PHP, Perl, Python, and XML.⁴⁶ Some languages are particularly well suited for business uses while others are more appropriate for capturing formal mathematical principles in research and academia.⁴⁷

37. *Id.* at 9.

38. *See, e.g.,* Campbell v. Facebook Inc., 310 F.R.D. 439, 448 (N.D. Cal. 2015) (describing how Facebook source code carries out processes).

39. Roig, *supra* note 14, at 327.

40. Foster & Northrop, *supra* note 35, at 43.

41. *Id.*

42. *Id.*

43. *Id.*

44. *Id.*

45. *Id.* at 44.

46. Roig, *supra* note 14, at 327; *see* Fox, *supra* note 15, at 877–78; Mislow, *supra* note 12, at 39. The Mislow article was written over three decades ago and it remains one of the clearest expositions in print of the programming process.

47. Fox, *supra* note 15, at 877–78.

The programmers begin writing the source code in the selected language.⁴⁸ A programmer familiar with the particular language can both read and write source code.⁴⁹ The source code itself is a combination of words and mathematical symbols⁵⁰ that have a particular meaning in the selected language.⁵¹ The original programmer writes the program logic into the code,⁵² and other programmers literate in the same language can follow the complete logical pathway of the program at a later time.⁵³

At this point in the process, the programmers should test the source code. Although the code has not yet been finalized and converted into machine code, the codewriters can use specialized development tools such as a “debugger” to conduct some limited testing to determine whether the instructions embedded in the source code suffice to accomplish the specified tasks.⁵⁴ If the source code passes muster in the testing, it is ready to be converted into machine or object code.

The source code must be distinguished from the machine or object code.⁵⁵ The latter code is binary⁵⁶—a series of 1’s and 0’s⁵⁷—that would strike most persons, including most programmers, as meaningless.⁵⁸ To be sure, there are some “super geeks” who can read and write binary code.⁵⁹ To write machine code, a programmer must employ a high-level assembly language rather than one of the languages previously mentioned.⁶⁰

The manual writing of machine code is both tedious and error prone.⁶¹ Consequently, in most cases, the programmers resort to a device called a compiler⁶² to convert⁶³ or transform⁶⁴ the source code

48. *People v. Robinson*, 860 N.Y.S.2d 159, 163 (App. Div. 2008) (describing source code as text “written onto a computer chip”).

49. Foster & Northrop, *supra* note 35, at 43; Roig, *supra* note 14, at 327.

50. Fox, *supra* note 15, at 877; see GRIMMELMANN, *supra* note 34, at 24–25 (providing examples of simple source code in Python and C, two different programming languages).

51. GRIMMELMANN, *supra* note 34, at 24–25.

52. Foster & Northrop, *supra* note 35, at 45; see Mislow, *supra* note 12, at 46–47.

53. Foster & Northrop, *supra* note 35, at 45.

54. *Id.* at 43.

55. See GRIMMELMAN, *supra* note 34, at 24–26.

56. Mislow, *supra* note 12, at 39.

57. *Id.*

58. Fox, *supra* note 15, at 880. Some commentators have asserted that machine language is virtually unreadable. Pamela Samuelson, *CONTU Revisited: The Case Against Copyright Protection for Computer Programs in Machine-Readable Form*, 1984 DUKE L.J. 663, 679 (1984) (analogizing a computer operating system to a secretary’s familiarity with an office filing system).

59. Foster & Northrop, *supra* note 35, at 43.

60. Fox, *supra* note 15, at 880.

61. *Id.*

62. *Id.*

into binary machine code.⁶⁵ In effect, the compiler translates the source code into machine or object code.⁶⁶ After that step, the source code's instructions are in a form that is directly executable by a computer.⁶⁷ The machine code can be run directly through the computer's processor in order to perform the tasks initially identified by the marketing personnel.⁶⁸

It should be evident why defendants challenging a software program prefer access to the source code as opposed to the machine or object code. Like the machine code, the source code contains all the instructions embedded in the program.⁶⁹ Unlike machine code, the source code is human readable by many programmers.⁷⁰ A source code is much more likely to be intelligible to a defense expert.⁷¹ Furthermore, there are commercially available tools such as Grep and Understand 2.0 that the programmer can use in evaluating source code.⁷² Defendants are much more likely to find an expert capable of critically evaluating that type of computer code if they can gain access to the source code.⁷³

III. CURRENT STATE OF THE LAW GOVERNING ADMISSIBILITY OF TESTIMONY BASED ON COMPUTERIZED FORENSIC SCIENCE TECHNIQUES AND THE DISCOVERABILITY OF THE COMPUTER PROGRAM SOURCE CODE

Most criminal cases raising source code issues have involved either breath testing devices or DNA probabilistic genotyping programs such as TrueAllele. A handful of criminal cases have addressed source code issues in other settings,⁷⁴ but those cases are few and far be-

63. Foster & Northrop, *supra* note 34, at 43.

64. Roig, *supra* note 14, at 327 (discussing how source code translated from readable form into object code).

65. Mislow, *supra* note 12, at 39.

66. Roig, *supra* note 14, at 327.

67. Fox, *supra* note 15, at 873 (discussing computer instructions); *Object Code*, TECHNOPEDIA, <https://technopedia.com/definition/546/object-code> (last visited Aug. 9, 2016).

68. Fox, *supra* note 15, at 880.

69. Roig, *supra* note 14, at 327.

70. Foster & Northrop, *supra* note 35, at 43.

71. *Id.* at 45.

72. *Id.*

73. See, e.g., *People v. Robinson*, 860 N.Y.S.2d 159, 167 (App. Div. 2008) (stating that the defendant failed to present evidence that the source code could be downloaded and analyzed from the device in an intelligible language).

74. See, e.g., *United States v. Chiaradio*, 684 F.3d 265, 271 (1st Cir. 2012) (discussing the F.B.I.'s use of the EP2P enhanced peer-to-peer file-sharing program in a child pornography prosecution); *United States v. Matish*, No. 16-cv-162016 U.S. Dist. LEXIS 82279 (E.D. Va. June 23, 2016).

tween. For that reason, this Part concentrates on breath-testing-device decisions and the TrueAllele case law.

A. The Admissibility of Prosecution Testimony Based on Automated Forensic Techniques

This Part reviews the legal framework to determine the admissibility of prosecution expert testimony based on the infrared breathing and the TrueAllele techniques.

1. Infrared Breath Testing Techniques

There have been two major waves of case law, the first dealing with testimony about infrared breath testing devices in drunk driving cases. In some jurisdictions, when the prosecution offers such testimony, the prosecution can invoke a statute that purports to eliminate the need for foundational testimony establishing the reliability of the technique.⁷⁵ Minnesota and New York are illustrative.

In Minnesota, prosecutors had the benefit of the following state statute:

In any civil or criminal hearing or trial, the results of a breath test, when performed by a person who has been fully trained in the use of an infrared or other approved breath-testing instrument, . . . pursuant to training given or approved by the commissioner of public safety or the commissioner's acting agent, are admissible in evidence without antecedent expert testimony that an infrared or other approved breath-testing instrument provides a trustworthy and reliable measure of the alcohol in the breath.⁷⁶

Even though courts are the government bodies that use evidentiary rules on a daily basis, state constitutions of most jurisdictions give the legislature the plenary power to prescribe evidentiary rules for the courts.⁷⁷ Hence, while the prosecution must ordinarily lay a foundation showing the reliability of a scientific technique employed by one of its witnesses,⁷⁸ the legislature can provide that in specified situations the prosecution need not present such testimony. In the words of the Minnesota statute, the prosecution may introduce testimony about the result of a breath testing device “without antecedent expert

75. *In re Comm'r of Public Safety*, 735 N.W.2d 706, 710 (Minn. 2007).

76. MINN. STAT. § 634.16 (2015).

77. Paul C. Giannelli, *The Proposed Ohio Rules of Evidence: The General Assembly, Evidence, and Rulemaking*, 29 CASE WES. RES. L. REV. 16, 24–25 (1978). *But see* *Brim v. State*, 624 N.E.2d 27, 33 n.2 (Ind. Ct. App. 1993) (striking down an evidentiary statute because the statute “conflict[ed] with the rule adopted in” an Indiana Supreme Court decision).

78. *See* 1 GIANNELLI ET AL., *supra* note 19, § 1.12 (describing the process for laying the foundation for a proper scientific technique).

testimony that” the instrument is “reliable” so long as: (1) the person who conducted the test complied with the administrative training requirements; and (2) that person used a testing device approved by the Commissioner of Public Safety.⁷⁹ Pursuant to the statute, the Minnesota Commissioner of Public Safety promulgated a list of approved breath testing devices, which are presumed reliable.⁸⁰ A defendant must rebut this presumption by establishing “reason to doubt the accuracy and reliability of the test results.”⁸¹ The published Minnesota opinions involve a product manufactured by a Kentucky corporation CMI, Inc., the Intoxilyzer 5000EN, an infrared device on Minnesota’s approved list.⁸²

The state of the law in New York is very similar. Like Minnesota, New York maintains a list of approved breath testing devices, developed by the New York State Department of Health.⁸³ At the time of the New York litigation, the list included a reference to “the Intoxilyzer 5000.”⁸⁴ However, as the New York courts acknowledged, CMI has released a series of 5000 breath testing devices.⁸⁵ The courts construed the statute to mean that testimony about any 5000 series device was admissible “without expert testimony” establishing the reliability of the device.⁸⁶ As in Minnesota, any listed device is “presumed reliable;”⁸⁷ and a defendant has the burden of showing that “software changes and upgrades” in later versions of the 5000 series device rendered the modified version of the device “unreliable.”⁸⁸

The Minnesota and New York cases made short shrift of defendants’ attempts to raise doubts about the reliability of the testing devices. The Minnesota courts rejected the defendants’ attempts as

79. MINN. STAT. § 634.16.

80. *In re Comm’r of Public Safety*, 735 N.W.2d 706, 710 (Minn. 2007); *State v. Underdahl*, 749 N.W.2d 117, 120–21 (Minn. Ct. App. 2008), *rev’d in part, aff’d in part*, 767 N.W.2d 677 (Minn. 2009) (affirming one appellant’s order to produce source code, while reversing another appellant’s order denying the production of source code).

81. *Underdahl*, 749 N.W.2d at 121.

82. *Id.* at 120; *In re Comm’r of Public Safety*, 735 N.W.2d. at 710.

83. N.Y. COMP. CODES R. & REGS. tit. 10, § 59.4 (2010).

84. *Cialino*, 831 N.Y.S.2d at 681.

85. *Id.*

86. *Id.* at 682.

87. *Id.*

88. *Id.* But see *Commonwealth v. Camblin*, 31 N.E.3d 1102 (Mass. 2015). In *Camblin*, the court refused to apply the Massachusetts statute analogous to the Minnesota and New York statutes. *Id.* at 1112–13. The court dealt with the Alcotest 7110 MK III-C, manufactured by Draeger Safety Diagnostics, Inc. *Id.* at 1104. The device was a “new generation,” dual sensoric instrument employing a fuel cell test as well as an infrared test. *Id.* at 1108, 1110. The court held that since the instrument was the “first” device of its kind, the statute did not automatically insulate it from a reliability challenge. *Id.* at 1112.

“inadequate.”⁸⁹ The New York cases are even more dismissive; they characterized the defense arguments as mere “fishing expedition[s].”⁹⁰ The decisions place the burden squarely on the shoulders of the defense.⁹¹ Without expert testimony identifying faults in the source code for the Intoxilyzer 5000’s software, defendants have been unable to meet that burden.

2. *The TrueAllele Technique*

In the leading Intoxilyzer 5000 cases in Minnesota and New York, the prosecution had the benefit of a state statute obviating the need for foundational expert testimony so long as the prosecution witness used a breath testing device on the state’s approved product list. To date, no jurisdiction has enacted a statute dispensing with the need for foundational testimony when the prosecution proffers testimony about a TrueAllele analysis of the genetic profiles of the contributors to a mixed DNA sample. Even without the benefit of such statutes, prosecutors have generally succeeded in introducing testimony based on the TrueAllele technique. The testimony has been admitted over defendants’ objections in at least seven states.⁹² Rather than invoking a statute eliminating the need for expert testimony, prosecutors in the TrueAllele cases have followed the conventional approach of laying a foundation by presenting testimony about validation studies establishing the reliability of the scientific technique.⁹³ There have been sev-

89. *State v. Underdahl*, 749 N.W.2d 171 (Minn. Ct. App. 2008), *rev’d in part, aff’d in part*, 767 N.W.2d 677 (Minn. 2009).

90. *People v. Robinson*, 860 N.Y.S.2d 159, 166 (App. Div. 2008); *Cialino*, 831 N.Y.S.2d at 682.

91. *Cialino*, 831 N.Y.S.2d at 682.

92. Palazzolo, *supra* note 27; see DAVID KAYE ET AL., *THE NEW WIGMORE, A TREATISE ON EVIDENCE: EXPERT EVIDENCE* 263–77 (2016 Supp.); Moss, *supra* note 23, at 1061–70. Moss discusses the following decisions:

- *Queen v. Colin Duffy & Brian Shivers*, [2011] NICC 37 (N. Ir. Crim.). This case was a prosecution in Northern Ireland. The judge conceded that as of 2011, only a “very small portion” of the world’s DNA laboratories used TrueAllele. Moss, *supra* note 23, at 1065. The judge relied heavily on a recommendation by the New York Commission on Forensic Science DNA Subcommittee that New York adopt the technique. *Id.* at 1065.
- *Commonwealth v. Foley*, 38 A.3d 882 (Pa. Super. Ct. 2012). The court ruled that the available validation studies established the reliability of TrueAllele. Moss, *supra* note 23, at 1067.
- *Ohio v. Shaw*, CR-13-575691, at *21 (Cuyahoga Ct. C.P. Oct. 10, 2014). The court noted that other jurisdictions had admitted testimony based on TrueAllele and added that the validation studies constituted sufficient testing of the technique. Moss, *supra* note 23, at 1070.

Id.

93. Moss, *supra* note 23, at 1061–70. There have also been tests of STRmix, another probabilistic genotyping program. *People v. Muhammad*, No. 14-56263-FC, at *5 (Mich. Cir. Ct. Dec. 15, 2015), http://media.mlive.com/chronicle/news_impact/other/muhammad-opinion.pdf; Josh S. Hausman, *Michigan Judge’s Landmark DNA Ruling Could Revolutionize CSI Work*, *MLIVE*

eral published validation studies for TrueAllele.⁹⁴ In a statement on the Cybergenetics website, the developer of TrueAllele, Dr. Mark Perlin, asserted that courts admitting TrueAllele testimony have reached the right result because “[s]cientists test executable software programs on real data [in validation studies]; they do not read source code text.”⁹⁵

*B. The Discoverability of the Source Code for the Software
Automated Forensic Techniques*

Courts have not only had to grapple with questions of the admissibility of automated forensic techniques without foundational testimony about the source code of the software controlling the automation, but they have also been forced to address the pretrial discoverability of the source code. While the prosecution has almost uniformly prevailed in cases analyzing the admissibility of the trial

(Dec. 23, 2015, 12:03 PM), http://www.mlive.com/news/muskegon/index.ssf/2015/12/michigan_judges_landmark_dna_r.html.

94. Paula Reed Ward, *Judge to Allow DNA Evidence in Deaths of Wolfe Sisters*, PITT. POST-GAZETTE (Oct. 30, 2015, 12:00 AM), <http://www.post-gazette.com/local/city/2015/10/30/allow-dna-evidence-in-deaths-of-East-Liberty-Wolfe-sisters/stories/201510300182>; see Susan A. Greenspoon et al., *Establishing the Limits of TrueAllele Casework: A Validation Study*, 60 J. FORENSIC SCI. 1263 (2015); Mark W. Perlin et al., *TrueAllele Casework on Virginia DNA Mixture Evidence: Computer and Manual Interpretation in 72 Reported Criminal Cases*, 9 PLoS ONE (2014); Mark W. Perlin et al., *New York State TrueAllele Casework Validation Study*, 58 J. FORENSIC SCI. 1458 (2013); Mark W. Perlin et al., *TrueAllele Genotype Identification on DNA Mixtures Containing up to Five Unknown Contributors*, 60 J. FORENSIC SCI. 857 (2015); Mark W. Perlin et al., *Validating TrueAllele DNA Mixture Interpretation*, 56 J. FORENSIC SCI. 1430 (2011); see also *Case Study: TrueAllele Validation on 16 Case Mixture Items*, THE DNA INVESTIGATOR, Spring 2010, at 2; Steven Myers & Jeanette Wallin, *STRmix V2.0.6 BFS Casework Internal Validation Summaries*, EPIC (2016), <https://epic.org/state-policy/foia/dna-software/EPIC-16-02-02-CalDOJ-FOIA-20160219-STRmix-V2.0.6-Validation-Summaries.pdf> (study evaluating the use of TrueAllele by the California Department of Justice Laboratory in Richmond, California). In a presentation entitled *Fighting for DNA Justice: Genotyping Software in the Hillary Acquittal*, in October 2016 in New York, Dr. Mark Perlin stated that there have been thirty-four validation studies of TrueAllele, including seven published studies. See Mark W. Perlin, *Fighting for DNA Justice: Genotyping in the Hillary Acquittal*, CYBERGENETICS (Oct. 2016), <https://www.cyngen.com/information/presentations/2016/Legal-Aid/Perlin-Fighting-for-DNA-justice-genotyping-software-in-the-Hillary-acquittal/handout.pdf>. He added that on ten occasions, the technique has survived admissibility challenges under *Frye* or *Daubert*. *Id.* at 4. The slides for Dr. Perlin’s presentation are available from Cybergenetics. *Id.* at 1.

95. Seth Augenstein, *Access Denied: Source Code for DNA Software Remains Protected in Pa. Murder Trial*, FORENSIC MAG. (Feb. 5, 2016), <http://www.forensicmag.com/article/2016/02/access-denied-source-code-dna-software-remains-protected-pa-murder-trial>; see *Commonwealth v. Robinson*, No. CC 201307777 (Pa. Ct. C.P. Allegheny Cty. Feb. 4, 2016); see also Palazzolo, *supra* note 27 (discussing a recent ruling by Pennsylvania Judge Jill Rangos in the prosecution of Michael Robinson for homicide); Silver, *supra* note 27. In *Commonwealth v. Foley*, the court stated that “scientists can validate the reliability of a computerized process even if the source code underlying that process is not available to the public.” 38 A.3d 882, 889 (Pa. Super. Ct. 2012).

testimony, the results on the pretrial discoverability issue have been more mixed. Once again the cases have come in two waves.

1. *Infrared Breath Testing Techniques*

Courts are divided over whether the defendants have a pretrial discovery right to access to the source code for breath testing devices. In three jurisdictions—Connecticut,⁹⁶ Florida,⁹⁷ and New York⁹⁸—the court rebuffed the defendants' efforts to gain access to the source code for an evaluation by defense experts. The courts reasoned that under its contracts with the manufacturer, the government had neither ownership nor actual or constructive possession of the code.⁹⁹ In addition, the courts concluded that the manufacturer could not be forced to divulge the code because the source code constitutes a protected trade secret.¹⁰⁰ Every American jurisdiction recognizes an evidentiary privilege protecting trade secrets.¹⁰¹

In contrast, Minnesota courts have permitted discovery after carefully reviewing the language of the contract between the government and CMI. Initially, an intermediate Minnesota appellate court read the contract and found that the source code was "not proprietary" and that the source code for the model sold to Minnesota was "created specifically for the Minnesota model of the Intoxilyzer."¹⁰² Later, the Minnesota Supreme Court ruled that under the terms of the state's original request for proposal, the code was the state's property.¹⁰³ The trial court had ordered disclosure of the code and ruled that un-

96. *See* *State v. Burnell*, No. MV06479034S, 2007 WL 241230 (Conn. Super. Ct. Jan. 18, 2007); *State v. Walters*, No. DBDMV050340997S, 2006 WL 785393 (Conn. Super. Ct. Feb. 15, 2006).

97. *See* *Moe v. State*, 944 So. 2d 1096 (Fla. Dist. Ct. App. 2006).

98. *See* *People v. Robinson*, 860 N.Y.S.2d 159, 168 (App. Div. 2008); *Cialino*, 831 N.Y.S.2d at 681.

99. *See, e.g.,* *People v. Cialino*, 831 N.Y.S.2d 680, 681 (Crim. Ct. 2007).

100. *Burnell*, 2007 WL 241230, at *2 ("CMI, Inc., manufacturer of the Intoxilyzer 5000, through its counsel, has declined to produce the items, claiming they constitute trade secrets."); *Moe*, 944 So. 2d at 1097 ("It is . . . without dispute that the code is a trade secret of CMI, Inc. and that CMI, Inc. has invoked its statutory and common law privileges protecting the code from disclosure. Therefore, the State cannot obtain possession of the code."); *Robinson*, 860 N.Y.S.2d at 167 ("The Intoxilyzer source code was not the property of the State, since it was owned and copyrighted by its manufacturer, CMI, Inc., a Kentucky corporation, and is a trade secret of CMI, Inc."); *Cialino*, 831 N.Y.S.2d at 681 (holding that the source code was "the property of a corporation that invoked statutory and common law privileges protecting the code from disclosure, thereby making it unobtainable").

101. 2 EDWARD J. IMWINKELRIED, *THE NEW WIGMORE, A TREATISE ON EVIDENCE: EVIDENTIARY PRIVILEGES* § 9.2.1 (2d ed. 2010).

102. *In re Comm'r of Public Safety*, 735 N.W.2d 706, 709–10 (Minn. 2007).

103. *State v. Underdahl*, 767 N.W.2d 677, 686 (Minn. 2009). The court noted that "[o]ne provision of the RFP, titled 'ownership of copyright,' states that any copyrightable material would 'be the property of the State and are by this Contract assigned to the State.'" *Id.* at 686 n.6.

less the state complied with the order, testimony about the Intoxilyzer test result would be excluded and the charge dismissed.¹⁰⁴ The Minnesota Supreme Court affirmed the order for the production of the complete computer source code.¹⁰⁵

2. *The TrueAllele Technique*

As we have seen, courts in the first wave of cases were divided over the discoverability of the source code for infrared breath testing devices. However, even then the prevailing view was that defendants were not entitled to discovery. In the second wave of cases, in which the focus shifted to probabilistic genotyping programs, notably TrueAllele, there has been a greater consensus among the courts. To date, although the issue has been litigated in at least seven states, no state court has ordered discovery of the TrueAllele source code.¹⁰⁶ Rather, harking back to precedent from Connecticut, Florida, and New York, the courts have uniformly denied discovery because the source code constitutes a trade secret and Cybergenetics has properly asserted the evidentiary privilege protecting trade secrets.¹⁰⁷

IV. A CRITICAL EVALUATION OF THE CURRENT STATE OF THE LAW

This Part addresses whether prosecution should be permitted to introduce expert testimony based on computerized techniques without presenting foundational testimony about the validity of the source code. This Part concludes that prosecution may do so; however, there are circumstances in which defendants should have access to the source code in order to challenge the weight of the testimony about the technique.

104. *Id.* at 680.

105. *Id.* at 687; *see also* *State v. Chun*, 943 A.2d 114, 174 (N.J. 2008) (requiring a breathalyzer manufacturer to disclose any future changes to software code). In an earlier decision, 923 A.2d 226, 226–27 (N.J. 2007), the court granted the defense discovery motion but stated only that “good cause appear[ed]” for disclosure.

106. *See, e.g.*, *Commonwealth v. Robinson*, No. CC 201307777 (Pa. Ct. C.P. Feb. 4, 2016); *see also* *Palazzo, supra* note 27; *Silver, supra* note 27. In one case, a California court ordered discovery, but the intermediate appellate court reversed the order, and the California Supreme Court declined to review the reversal. *People v. Superior Court (Chubbs)*, No. B258569, 2015 WL 139069 (Cal. Ct. App. Jan. 9, 2015). In *United States v. Kevin Johnson*, a federal District Court ordered discovery of the source code for the Forensic Statistical Tool used by the New York City Office of the Chief Medical Examiner. No. 15-CR-565 (VEC) (S.D.N.Y. June 7, 2016). In her order, Judge Valerie Caproni stated that she was “prepared to enter a protective order if OCME wishes” *Id.* at 1.

107. *Robinson*, 860 N.Y.S.2d at 167 (holding that the software source code would be not be admitted because its release would cause irreparable harm to the manufacturer); *Palazzo, supra* note 27.

*A. The Admissibility of Prosecution Testimony Based on
Automated Forensic Techniques*

The threshold question is whether the prosecution may introduce evidence based on an automated forensic technique, without presenting testimony validating source code of the software controlling the automation. To answer that question, we must distinguish two types of jurisdictions.

1. Jurisdictions That Do Not Require a Showing of the Scientific Validation of a Forensic Technique as Part of the Foundation for Admissibility

To begin, suppose the legislature has enacted a statute that, if fairly interpreted, manifests a legislative intent to permit the introduction of evidence, based on a forensic technique, without the need for foundational testimony about the scientific validation of the technique. In most jurisdictions the state legislature has the power to prescribe evidentiary rules for the courts.¹⁰⁸ Some state legislatures have exercised power to provide statutorily for the admission of testimony based on forensic techniques such as breath testing devices¹⁰⁹ and DNA typing.¹¹⁰ Unless the court invalidates the statute on some constitutional ground, the statute eliminates the need for the prosecution to present any foundational testimony about the empirical validity of the technique. In these jurisdictions, there is no need for the prosecution to show the validity of the source code in the software for the automated technique.

In addition, although the majority of states¹¹¹ have now adopted some variation of the empirical validation standard that the U.S. Su-

108. Giannelli, *The Proposed Ohio Rules of Evidence*, *supra* note 77, at 20 (discussing the factors the Ohio General Assembly uses when promulgating the Ohio Rules of Evidence). *But see* *Brim v. State*, 624 N.E.2d 27, 33 n.2 (Ind. Ct. App. 1993) (striking down an evidentiary rule because the statute “conflict[ed] with the rule adopted in” an Indiana Supreme Court decision).

109. *See, e.g.*, MINN. STAT. § 634.16 (2015); N.Y. VEH. & TRAF. § 1194(b) (McKinney 2010) (stating that a breath test may be administered by a police officer).

110. *See, e.g.*, IND. CODE § 35-37-4-13(b) (“In a criminal trial or hearing, the results of forensic DNA analysis are admissible in evidence without antecedent expert testimony that forensic DNA analysis provides a trustworthy and reliable method of identifying characteristics in an individual’s genetic material.”); Elizabeth Marie Bezak, Note, *DNA Profiling Evidence: The Need for a Uniform and Workable Evidentiary Standard of Admissibility*, 26 VA. U.L. REV. 595, 620 nn.175–79 (1992) (listing state statutes); Jayne L. Jakubaitis, Note, ‘Genetically’ Altered Admissibility: Legislative Notice of DNA Typing, 39 CLEV. ST. L. REV. 415, 426 (1991); Note, *A Defense Counsel’s Guide to the Use of DNA Evidence in Virginia*, 2 REGENT U.L. REV. 113, 114 n.6 (1992) (listing state statutes).

111. 1 GIANNELLI ET AL., *supra* note 19, §§ 1.14–1.15 (noting a total of thirty states have moved to *Daubert* analysis of evidence).

preme Court announced in its 1993 decision, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,¹¹² a minority of states still adhere to the traditional *Frye*¹¹³ “general acceptance” test. Under the *Frye* test, the proponent of scientific testimony need not show the empirical validation of a scientific technique or theory.¹¹⁴ Rather, the proponent must demonstrate that the technique or theory has gained a certain degree of popularity—“general acceptance”—within the relevant scientific fields.¹¹⁵ Courts favoring the general acceptance standard have doubts about the competence of lay jurors and trial judges to critically evaluate scientific testimony.¹¹⁶ Given those doubts, courts have embraced general acceptance as a proxy¹¹⁷ for the scientific validity of the theory or technique. In effect, these courts delegate the admissibility decision to the scientific community. Unlike the states with statutes authorizing the admission of particular types of scientific evidence without expert foundational testimony, these jurisdictions require such testimony. This required testimony must speak to the extent of the consensus supporting the technique rather than the underlying empirical validation.

2. Jurisdictions That Require a Showing of the Scientific Validation as Part of the Foundation for Admissibility

Until the early 1970s, federal courts and the overwhelming majority of states used *Frye*’s general acceptance standard as the litmus test for determining the admissibility of scientific testimony.¹¹⁸ At one point, *Frye* was the law in federal court and also in forty five states;¹¹⁹ however, in 1975, the Federal Rules of Evidence took effect.¹²⁰ Article VII of the new rules dealt with expert opinion testimony.¹²¹ Article VII did not contain any language that could reasonably be interpreted as codifying the traditional general acceptance test. Moreover, Fed-

112. 509 U.S. 579 (1993).

113. *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923), *superseded by rule as recognized in Daubert*, 509 U.S. 579.

114. *Daubert*, 509 U.S. at 590.

115. 1 GIANNELLI ET AL., *supra* note 19, § 1.06[c].

116. *Id.* § 1.06[a].

117. Bert Black et al., *Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge*, 72 TEX. L. REV. 715, 718–19 (1994); see Edward J. Imwinkelried, *The End of the Era of Proxies* 8 (UC Davis Legal Studies Research Paper Series, Working Paper No. 250, 2011), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1777551.

118. 1 GIANNELLI, ET AL., *supra* note 19, § 1.06.

119. Betty R. Steingass, *Changing the Standard for the Admissibility of Novel Scientific Evidence*: *State v. Williams*, 40 OHIO ST. L.J. 757, 769 n.83 (1979).

120. Act of Jan. 2, 1975, Pub. L. No. 93-595, 88 Stat. 1926 (codified at 28 U.S.C. §§ 2072, 2075 (2012)).

121. *Id.* art. VII, 88 Stat. 1937.

eral Rule 402 manifested a legislative intent to abolish uncodified restrictions on the admission of logically relevant evidence that satisfied the exclusionary rules set out in the rules.¹²² It was against this backdrop that the Supreme Court handed down its *Daubert* decision in 1993.¹²³

In *Daubert*, the Court initially ruled that the general acceptance test was no longer good law in federal practice.¹²⁴ Writing for the Court, Justice Blackmun relied on Rule 402¹²⁵ and stated that he could not find any language in the statutory text that could reasonably bear the interpretation that it incorporated the *Frye* test.¹²⁶ He wrote that the general acceptance standard was an “austere”¹²⁷ test at odds with the “liberal,”¹²⁸ “permissive”¹²⁹ thrust of the Federal Rules.

Justice Blackmun quickly added that the abolition of the general acceptance test did not mean expert testimony was freely admissible in federal court.¹³⁰ At this juncture in the opinion, Justice turned to the text of Rule 702 governing the admissibility of “scientific, technical, or other specialized knowledge.”¹³¹ The Justice focused on the interpretation of the expression, “scientific . . . knowledge.”¹³² In his mind, Congress’ choice of the noun “knowledge” meant that an expert opinion must rest on more than an expert’s subjective belief or unsubstantiated speculation.¹³³ The crucial question then became the meaning of the adjective “scientific.” Drawing on a number of amicus briefs filed by scientists and scientific organizations,¹³⁴ Justice Blackmun defined the adjective in methodological terms.¹³⁵ He referred to

122. Edward J. Imwinkelried, *Federal Rule of Evidence 402: The Second Revolution*, 6 REV. LITIG. 129, 139 (1987). At the time of its enactment, Rule 402 stated that all logically relevant evidence is admissible unless the judge can justify excluding the evidence under the Constitution, another provision of the Federal Rules, or other rules adopted by the Supreme Court pursuant to statutory authority (i.e. the Federal Rules of Civil and Criminal Procedure). *Id.* at 130. *Frye* was a creature of case law. Construed in light of Rule 1001(b)(5), restyled Rule 402 is to the same effect.

123. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 587 (1993).

124. *Id.* at 597–98.

125. *Id.* at 588.

126. *Id.* at 589.

127. *Id.*

128. *Id.* at 588; see *Almonte v. Averna Vision & Robotics, Inc.*, 128 F. Supp. 3d 729, 738 (W.D.N.Y. 2015).

129. *Daubert*, 509 U.S. at 589.

130. *Id.*

131. FED. R. EVID. 702.

132. *Daubert*, 509 U.S. at 590 (quoting FED. R. EVID. 702).

133. *Id.*

134. *Id.* at 581.

135. Edward J. Imwinkelried, *The Daubert Decision: Frye Is Dead, Long Live the Federal Rules of Evidence*, TRIAL, Sept. 1993, at 60, 62–63.

the classic Newtonian methodology of formulating a hypothesis, subjecting the hypothesis to empirical testing to falsify or validate the hypothesis, and lastly evaluating the results of the testing.¹³⁶ Again relying on the amicus briefs, he acknowledged the modern understanding that even when a scientist properly uses this methodology, he or she cannot achieve absolute certainty.¹³⁷ The Justice explained that to qualify as sufficiently reliable and admissible “scientific . . . knowledge,” a technique or theory must rest on adequate validation.¹³⁸ The Court declared that trial judges have a “gatekeeping”¹³⁹ responsibility to ensure that proffered scientific testimony passes muster. To help trial judges discharge that responsibility, Justice Blackmun listed several factors that trial judges could consider, including whether the hypothesis is empirically testable, whether it has been subjected to testing, whether the error rate was ascertained from the testing, and whether there are recognized procedures for conducting the testing.¹⁴⁰

Daubert was the first in a trilogy of high-profile Supreme Court decisions on the admissibility of the expert testimony. After rendering *Daubert* in 1993, the Court handed down *General Electric Co. v. Joiner*¹⁴¹ in 1997 and *Kumho Tire Co. v. Carmichael*¹⁴² in 1999. Collectively, the Court’s teaching is that the proponent of expert must make a foundational showing that by using the particular technique or theory¹⁴³ the expert employed, the expert can accurately draw the

136. *Daubert*, 509 U.S. at 590.

137. *Id.* (“[A]rguably, there are no certainties in science.”); see Edward Imwinkelried, *The Importance of Forensic Metrology in Preventing Miscarriages of Justice: Intellectual Honesty About the Uncertainty of Measurement in Scientific Analysis*, 7 J. MARSHALL L.J. 333, 346–47 (2014).

If the outcomes of numerous empirical tests of the hypothesis all confirm the hypothesis, we can have increasing confidence in the validity of the hypothesis. However, we cannot regard the hypothesis as “definitively confirmed because it is always possible that an empirical test will someday demonstrate the theory to be incorrect.” Another empirical test is always conceivable; and so long as that is true, a theoretical possibility of falsification or disproof remains When an expert relies on inductive reasoning to investigate the truth of a hypothesis, at most, the hypothesis can be accepted contingently or provisionally.

Id.; see Edward J. Imwinkelried, *Evidence Law Visits Jurassic Park: The Far-Reaching Implications of the Daubert Court’s Recognition of the Scientific Enterprise*, 81 IOWA L. REV. 55, 58 (1995).

138. *Daubert*, 509 U.S. at 589–90 (quoting FED. R. EVID. 702).

139. *Id.* at 597.

140. *Id.* at 593–94.

141. 522 U.S. 136 (1997).

142. 526 U.S. 137 (1999).

143. D. Michael Risinger, *Defining the “Task at Hand”: Non-Science Forensic Science After Kumho Tire Co. v. Carmichael*, 57 WASH. & LEE L. REV. 767, 773 (2000) (noting that the Supreme Court’s trilogy of decisions, *Daubert*, *Joiner*, and *Kumho* makes it clear that the question is not the “global validity” of the discipline or field); see Edward J. Imwinkelried, *The Meaning*

specific type of inference¹⁴⁴ that he or she contemplates testifying to. The Court's reliability test has gradually garnered support among the states, and it is now the dominant test used by a majority of states.¹⁴⁵

We must ask: In a *Daubert* jurisdiction, must the proponent of evidence based on a computerized forensic technique present expert testimony about the computer's source code as part of the foundation for the evidence? Even in a *Daubert* jurisdiction requiring proof of validation as an essential part of the foundation, the answer is "no." The proponent can meet that burden without proffering testimony about the validity of the source code. Instead, as prosecutors have successfully argued in the TrueAllele cases,¹⁴⁶ the government can meet that burden by presenting testimony about validation studies investigating the accuracy of the software.

There is often more than one way that the proponent of expert testimony can lay an adequate foundation validating the underlying theory or technique. For instance, suppose that the hypothesis relates to a proposition in the field of toxicology, such as the theory that exposure to a particular pesticide can cause a certain type of cancer. The proponent of that theory can rely on formal epidemiological studies to validate that hypothesis.¹⁴⁷ Absent epidemiological data, some jurisdictions accept a combination of animal studies and chemical structure analyses as an alternative method of validation.¹⁴⁸ On the other hand, suppose that the proponent attempts to introduce testimony based on a new DNA testing kit. The court may dispense with the need for proof of new validation studies for the novel kit if the proponent can show that with minor changes, the new kit represents a modest modification of a previously validated DNA typing methodology.¹⁴⁹

The very purpose of a validation study is to investigate whether the theory or technique does what its proponent claims.¹⁵⁰ Suppose that the issue is whether TrueAllele enables an expert to accurately deter-

of "Appropriate Validation" in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, *Interpreted in Light of the Broader Rationalist Tradition, Not the Narrow Scientific Tradition*, 30 FLA. ST. U.L. REV. 735, 740-42 (2003) (reviewing the language in the three cases that indicates that the analytic focus has narrowed to the specific technique or theory that the expert contemplates using).

144. See *Joiner*, 522 U.S. at 144 (discussing the admissibility of "these experts' opinions").

145. See 1 GIANNELLI ET AL., *supra* note 19 §§ 1.14-1.15.

146. Palazzolo, *supra* note 27.

147. 1 GIANNELLI ET AL., *supra* note 19, § 15.05[d] (discussing epidemiological studies as a method of proving causation).

148. *Id.*

149. See *People v. Cordova*, 358 P.3d 518, 538 (Cal. 2015).

150. See Edward J. Imwinkelried, *Coming to Grips with Scientific Research in Daubert's "Brave New World": The Courts' Need to Appreciate the Evidentiary Differences Between Validity and Proficiency Studies*, 61 BROOK. L. REV. 1247, 1256-60 (1995).

mine the genetic profiles of multiple contributors to a mixed DNA sample, and that in well-designed validation studies the use of TrueAllele enables researchers to do precisely that. Federal Rule of Evidence 901(b)(9) captures the essence of the “authentication” or validation of a scientific technique.¹⁵¹ In the words of 901(b)(9), the essential foundation is a “showing that [the process or system] produces an accurate result.”¹⁵² Validation studies summarizing the results of tests of the technique and showing that the technique yields accurate results satisfy that standard. As a matter of logic, the court should treat the studies as adequate validation under *Daubert*. The proponent can shoulder the burden under *Daubert* without making a further, separate showing about the source code of the software controlling TrueAllele. The lack of testimony about the source code might increase the degree of uncertainty in the expert’s final opinion, but post-*Daubert*, the expert need not vouch for his or her opinion as a certainty.¹⁵³ In short, many courts have reached the correct result that prosecution evidence based on TrueAllele can be admitted, even without testimony about the source code.¹⁵⁴

B. The Discoverability of the Source Code of the Software for the Automated Forensic Technique

Assume *arguendo* that a court rules the proponent has laid an adequate foundation for the admissibility of evidence, which is based on an automated forensic technique, without any showing of the accuracy of the software’s source code. Does the opponent have a pretrial discovery right to access the source code? Again, to date in the TrueAllele cases, all the state courts that have reached this question have denied defendants’ discovery requests.¹⁵⁵ Have the courts reached the correct result here, as they have on the admissibility question discussed in Subpart A? To resolve that question, we must assess both

151. FED. R. EVID. 901(b)(9).

152. *Id.*

153. *Daubert v. Merrell Dow Pharm. Inc.*, 509 U.S. 579, 590 (1993); see Imwinkelried, *The Importance of Forensic Metrology*, *supra* note 137, at 345–53 (discussing how given *Daubert*, the courts should neither require nor even permit experts to vouch for their opinions as absolute certainties).

154. See SHIRA A. SCHEINDLIN, ELECTRONIC DISCOVERY AND DIGITAL EVIDENCE IN A NUTSHELL 413–14 (2d ed. 2016) (discussing *United States v. Chiaradio*, 684 F.3d 265, 278 (1st Cir. 2012) (sustaining the admission of testimony based on the government enhanced peer-to-peer file-sharing program in a child pornography case even though the program’s source code had not been peer reviewed)).

155. Palazzolo, *supra* note 27; see *Commonwealth v. Robinson*, No. CC 201307777 (Pa. Ct. C.P. Allegheny Cty. Feb. 4, 2016); see also Silver, *supra* note 27.

the defendants' alleged need for access and the countervailing considerations.

1. *The Defense's Alleged Need for Access to TrueAllele's Source Code*

In the TrueAllele cases, defendants have urged that it is unfair to deny them access and treat the program's source code as a mysterious "black box."¹⁵⁶ Some may argue that it is a waste of time to grant the defense access to the source code if, *ex hypothesi*, testimony based on the automated forensic technique will prove nonetheless admissible in a jurisdiction following *Daubert*. This counterargument misses the mark. Even when the proponent's item of evidence is admissible, the opponent has the right to attack the weight or believability of the evidence.¹⁵⁷ Just as the jury need not believe all evidence, the opponent has a right to attack the weight of the proponent's admissible evidence. The U.S. Supreme Court has held that in criminal cases, the defendant's right to attack the weight of the prosecution's evidence is of constitutional dimension under the Sixth Amendment Confrontation Clause.¹⁵⁸

For that matter, it should come as no surprise that even if the judge rules an item of prosecution evidence admissible, there can be significant remaining questions about the weight and believability of the evidence. In *Daubert*, the Supreme Court announced that under Federal Rule of Evidence 702 the trial judge has an important "gatekeeping" responsibility to police the reliability of proffered expert testimony.¹⁵⁹ However, the Court made that announcement in an opinion in which it characterized the general spirit of the Federal Rules as "liberal"¹⁶⁰ and "permissive."¹⁶¹ Furthermore, Justice Blackmun emphasized that

156. Jessica Leber, *The Problem with Using Secret Computer Code to Put People in Jail*, CO.EXIST (Nov. 2, 2015, 8:01 AM), <http://www.fastcoexist.com/3052420/the-problem-with-using-secret-computer-code-to-put-people-in-jail>.

157. See *infra* notes 158–162 and accompanying text; see also Comment, *Pomona v. SQM North America Corp.: Trial Courts as "Gatekeeper" Under Daubert and its Progeny*, 39 AM. J. TRIAL ADVOC. 377, 387 (2015).

158. *Crane v. Kentucky*, 476 U.S. 683, 686–87 (1986); see *United States v. Bess*, 75 M.J. 70, 75–76 (C.A.A.F. Jan. 6, 2016) (ruling that certain muster reports are admissible at a hearing outside the jury's presence; the judge submitted the reports to the court-martial members without allowing the defendant to attack the weight of the reports in front of the members; the judge "does not limit the right of a party to introduce before the trier of fact evidence relevant to weight or credibility").

159. *Daubert v. Merrell Dow Pharm. Inc.*, 509 U.S. 579, 597 (1993).

160. *Id.* at 588; see *Seamon v. Remington Arms Co.*, 813 F.3d 983, 988 (11th Cir. 2016); *United States v. Rodriguez*, 125 F. Supp. 3d 1216, 1231–32 (D.N.M. 2015); *Veleron Holding, B.V. v. Stanley*, 117 F. Supp. 3d 404, 443 (S.D.N.Y. 2015).

161. *Daubert*, 509 U.S. at 589.

the focus of the judge's analysis must be an evaluation of the soundness of the expert's methodology in reaching his or her conclusion, not the conclusion itself.¹⁶² The Court's message has not been lost on either the Advisory Committee or the lower courts. In 2000, the Advisory Committee amended Federal Rule of Evidence 702, and in a note to the Amendment, the Committee approvingly quoted the Third Circuit Court of Appeals: Proponents are not required to "demonstrate to the judge by a preponderance of the evidence that the assessments of their experts are correct, they only have to demonstrate by a preponderance of the evidence that their opinions are reliable The evidentiary requirement of reliability is lower than the merits standards of correctness."¹⁶³ The lower courts have frequently stated that when evaluating the admissibility of the proffered expert testimony, their inquiry is "limited"¹⁶⁴ as well as "preliminary."¹⁶⁵ Indeed, the lower courts sometimes add that "[i]n general, Rule 702 has been interpreted to favor admissibility."¹⁶⁶ Construed together, these passages in *Daubert*, the lower court opinions, and the Advisory Committee Notes demonstrate the correctness of Justice Blackmun's observation in *Daubert* that evidence admissible under Rule 702 could still be "shaky."¹⁶⁷

How shaky could the weight of admissible testimony about an automated forensic technique be? A study of the source code for New Jersey's breath testing devices "uncovered a variety of defects that could impact the test results."¹⁶⁸ Yet, in the initial wave of source code decision involving breath testing devices, courts repeatedly asserted that there is little need for discovery of the source code because the defense had "other avenues of challenge" to the weight of the

162. *Id.*

163. FED. R. EVID. 702 advisory committee's note to 2000 amendment (quoting *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 744 (3d Cir. 1994), *cert. denied sub nom.* General Electric v. Ingram, 513 U.S. 1190 (1995)).

164. *DL v. Dist. of Columbia*, 109 F. Supp. 3d 12, 28 (D.D.C. 2015).

165. *In re NJOY, Inc. Consumer Class Action Litig.*, 120 F. Supp. 3d 1050, 1068–69 (C.D. Cal. 2015) (the Supreme Court itself used that adjective); *see Daubert*, 509 U.S. at 592.

166. *DL*, 109 F. Supp. 3d at 27–28 (quoting *Khairkhwa v. Obama*, 793 F. Supp. 2d 1, 10 (D.D.C. 2011)). The court also quotes the Advisory Committee's observation that "[a] review of the caselaw after *Daubert* shows that the rejection of expert testimony is the exception rather than the rule." FED. R. EVID. 702 advisory committee's note to the 2000 amendments.

167. *Daubert*, 509 U.S. at 596; *see Rodman v. Safeway Inc.*, 125 F. Supp. 3d 922, 937 (N.D. Cal. 2015); *Speaks v. Mazda Motor Corp.*, 118 F. Supp. 3d 1212, 1218 (D. Mont. 2015); *see also* Comment, *Pomona v. SQM North America Corp.: Trial Courts as "Gatekeeper" Under Daubert and Its Progeny*, 39 AM. J. TRIAL ADVOC. 377, 387 (2015) (noting that even if scientific testimony is admissible under *Daubert*, it can be "precarious").

168. *State v. Underdahl*, 767 N.W.2d 677, 685 (Minn. 2009) (describing a report submitted in *State v. Chun*, 943 A.2d 114, 132–33 (N.J. 2008)).

prosecution evidence.¹⁶⁹ At the same time that these courts denied the defendants access to the source code, they pointed out that the defendants had access to documentation, including the device's maintenance records and the checklist the analyst used in conducting the test.¹⁷⁰ Those records do not contain the same information that an examination of the software's source code would yield. The analyst's checklist might minimize the risk of human error in conducting a test at a specific time and place, but the checklist provides no insight into any inherent defects in the program logic. Likewise, maintenance records could prove that for a certain period after a maintenance the device was operating as intended; but again, even if the device was operating as intended, there might be a defect buried in the source code. In sum, the discoverability of those documents does not undercut the case for discovery of the source code.

The real issue is the soundness of Cybergenetics' contention that the availability of validation studies for TrueAllele eliminates any legitimate need to study the source code.¹⁷¹ In civil cases involving source code discovery, litigants have occasionally opposed discovery by arguing that the available validation studies enable the opponent to study "the correlation between input data and output data," so as to determine whether the program operates accurately.¹⁷² In the same vein, in a Minnesota case denying discovery of the source code for the Intoxilyzer 5000EN, the court concluded that a review of available validation studies allows an opponent to determine whether a computer program contains deficiencies "without access to the source code."¹⁷³

Is that conclusion correct? On close scrutiny, the answer is that the conclusion is correct—sometimes. The answer does not turn on the mere existence of validation studies or even their availability to the defense. Rather, the answer depends on the number of studies, their quality, the findings in the studies, and a comparison between the test conditions and the conditions in the instant case.

By way of example, assume that the trial judge construed Rule 702 so liberally that the judge allowed the prosecution to introduce testi-

169. *People v. Robinson*, 860 N.Y.S.2d 159, 165 (App. Div. 2008) (quoting *People v. Alvarez*, 515 N.E.2d 898, 900 (N.Y. 1987)).

170. *Robinson*, 860 N.Y.S.2d at 166 (discussing "calibration records," records which are "showing "that the machine was . . . properly maintained or that the test was . . . properly administered").

171. Palazzolo, *supra* note 27.

172. Mislow, *supra* note 12, at 46.

173. *State v. Underdahl*, 749 N.W.2d 117, 121–22, *aff'd in part, rev'd in part*, 767 N.W.2d 677 (Minn. 2009).

mony about the results of the use of a new automated forensic technique based on meager validation data—perhaps even a single validation study. This state of the record presents an appealing case for defense discovery. If the basis for admitting the testimony is so thin that it barely passes *Daubert* muster, the testimony would arguably remain “shaky;”¹⁷⁴ and defendants would have a powerful argument that the unresolved questions about the reliability of the technique cut in favor of pretrial discovery.

However, in the case of TrueAllele, there have been a number of published validation studies of the accuracy of the software.¹⁷⁵ Here the question becomes the extent or reach of the validation of the software. In *Joiner*,¹⁷⁶ the plaintiff relied on four studies, particularly studies that made use of animal research, to validate the hypothesis that exposure to a particular chemical was capable of causing the kind of cancer the plaintiff developed.¹⁷⁷ The plaintiff’s studies were undeniably suggestive of causation, but, in the Court’s mind, the question was the justifiability of extrapolating from those studies to the specific inference that the plaintiff’s expert proposed drawing.¹⁷⁸ The rub was that there were numerous differences between the published studies and the facts in *Joiner*.¹⁷⁹

174. See *Daubert v. Merrell Dow Pharm. Inc.*, 509 U.S. 579, 596 (1993). Especially if you posit the position of the President’s Council of Advisors on Science and Technology, it would be dubious to admit the testimony based on a single validation study. In its recent report, PRES. COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT—FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (2016), the Council discusses the concept of foundational validity. The Council takes the position that to be foundationally valid, a technique or theory must be supported by empirical research demonstrating that its results are both repeatable and reproducible. *Id.* at 47. “Repeatable” denotes “that, with known probability, an examiner obtains the same result, when analyzing samples from the same sources.” *Id.* at 79. In contrast, “reproducible” means “that, with known probability, different examiners obtain the same result, when analyzing the samples.” *Id.* It will often be difficult to treat a single study as an adequate demonstration of reproducibility.

175. See *supra* note 94 and accompanying text.

176. *General Electric Co. v. Joiner*, 522 U.S. 136, 146–47 (1997).

177. *Id.* at 144.

178. *Id.* at 146.

179. *Joiner*, 522 U.S. at 143–45 (discussing the differences between animal studies and the facts of the case). Although the plaintiff’s experts proposed opining about causation in human beings, the studies involved mice. *Id.* The Court did not say categorically that a proponent may never rely on animal studies, but the Court indicated that the judge should carefully scrutinize the parameters of an animal study before allowing an expert to treat the study as an adequate basis for an opinion about similar phenomena in human beings. *Id.* Precisely for that reason, the Court delved into other aspects of the animal studies. The animal studies involved infant mice while the plaintiff was an adult human being. *Id.* The chemicals were injected into the mice while the plaintiff had only dermal exposure. Given their respective sizes, the mice received “massive” doses compared to the limited exposure of the plaintiff. Some of the mice suffered from different illness than the cancer the plaintiff developed. *Id.*

The *Joiner* Court ultimately concluded that the trial judge had not abused her discretion in refusing to allow the plaintiff's experts to rest their opinions on the cited animal studies.¹⁸⁰ For that matter, in the seminal *Daubert* decision,¹⁸¹ the Court remarked that the expert evidence must "fit" the facts of the instant case—raising the question of the external validity of any study the expert relies on.¹⁸²

An analogy to metrology, the science of measurement, may be helpful.¹⁸³ The fundamental tenet of metrology is that to a degree, every measurement is uncertain.¹⁸⁴ It is impossible to know whether a measurement has accurately captured the true value of the measurand, the quantity or object being measured.¹⁸⁵ Metrology attempts to identify the sources of uncertainty in measurement and provide some sense of the extent of the uncertainty.¹⁸⁶ Metrologists have developed the notion of a "range of calibration."¹⁸⁷ According to the National Institute of Standards and Technology (NIST),

It is a generally accepted principle of reliable analysis that chemical analyzers should be calibrated over the full range of measurement and that measurement data be restricted to the range calibrated. It is not good measurement practice to report extrapolated data, i.e., outside of the range calibrated. The range of reliable calibration can be considered the range of reliable measurement and conversely.¹⁸⁸

NIST adds the caution that "[s]tandards should never be used in an extrapolative mode. They should always bracket the measurement range. No measurement should be reported at a value lower or higher than the lowest or highest standard used to calibrate the measurement process."¹⁸⁹ One Pennsylvania court endorsed this concept and held that readings outside the range of calibration (of the state's breath test machines) were unacceptable and inadmissible.¹⁹⁰

180. *Joiner*, 522 U.S. at 144.

181. *Daubert v. Merrell Dow Pharm. Inc.*, 509 U.S. 597, 589–91 (1993).

182. DAVID L. FAIGMAN ET AL., *SCIENCE IN THE LAW: STANDARDS, STATISTICS, AND RESEARCH ISSUES* § 8 (2002).

183. TED VOSK & ASHLEY R. EMERY, *FORENSIC METROLOGY: SCIENTIFIC MEASUREMENT AND INFERENCE FOR LAWYERS, JUDGES, AND CRIMINALISTS* § 1.4 (2015).

184. *Id.* § 6.2.

185. *Id.* § 1.4, Ch. 2.

186. *Id.*

187. *Id.* at 106.

188. *Id.* (quoting John K. Taylor, *STANDARD REFERENCE MATERIALS: HANDBOOK FOR SRM USERS* 7, U.S. DEPT. OF COMM. (1993)).

189. Taylor, *supra* note 188.

190. *Id.*; see *Commonwealth v. Schildt*, No. 2191 CR 2010, Opinion (Pa. Ct. C.P. Dauphin Cty. Dec. 31, 2012), <https://www.thenewspaper.com/rlc/docs/2013/pa-intoxilyzer.pdf>. For a discussion of *Commonwealth v. Schildt*, see Vosk & Emery, *supra* note 183, § 8.

In the context of this discovery issue, the courts should analogize to the notion of range of calibration and recognize a parallel notion of a range of validation. When scientists conduct a validation study, they control for certain variables and test the validity of the hypothesis under those conditions.¹⁹¹ A successful test or, better still, a series of successful tests supplies an objective basis for concluding that the hypothesis (the accuracy of a certain technique or theory) is valid under those conditions.¹⁹² However, test outcomes do not universally validate the use of the hypothesis under all conditions. Rather, the inference of the reliability of the technique or theory is fully warranted only when the controlled variables are the same.¹⁹³

During an admissibility analysis, most courts have not embraced such a rigid notion of range-of-validation to prohibit all extrapolation. It would be difficult to uphold such a prohibition given the Supreme Court's treatment of extrapolation in *Joiner*. The Court did not forbid extrapolation from empirical data by experts; instead, the Court cautioned that the trial judge must carefully compare the conditions in the studies with the conditions in the pending case to decide whether an extrapolation is defensible.¹⁹⁴ The *Joiner* Court's position is sensible. If in his or her work a scientist wants to know whether the same experimental result will obtain under different conditions, the scientist has the time and luxury of conducting another validation study under the changed conditions. However, by the time of trial, when an opponent points to a differing condition in the pending case, the failsafe point may have been passed, and it may be too late to conduct a further study before deciding whether to admit the proponent's testimony.

When the opponent can point to a difference in conditions that could plausibly lead to a different outcome, the proponent is either at the margins of the range of validation or has exceeded that range. Here, the opponent should presumptively be entitled to discovery to explore the significance of the suspected difference to determine how "shaky"¹⁹⁵ the difference renders the extrapolation. Pretrial discovery centered on such a difference would be far more than a speculative

191. Edward J. Imwinkelried, *The Debate in the DNA Cases Over the Foundation for the Admission of Scientific Evidence: The Importance of Human Error as a Cause of Forensic Misanalysis*, 69 WASH. U.L. Q. 19, 29 (1991).

192. *Id.*

193. *Id.*

194. *General Electric, Inc. v. Joiner*, 522 U.S. 136, 146 (1997) (describing conclusions extrapolated from data).

195. *Daubert v. Merrell Dow Pharm. Inc.*, 509 U.S. 579, 596 (1993).

fishing expedition.¹⁹⁶ If a validation study for a program such as TrueAllele involved only three contributors, but the pending prosecution is a gang rape prosecution alleging penetration by five contributors, there is a common sense question about the propriety of the extrapolation.¹⁹⁷ Likewise, if the validation study involved quantitatively equal fluid contributions but the analysis in the pending case found a great disparity among the volumes of the contributions by the various supposed rapists, questions would naturally arise from the difference in mixture proportions.¹⁹⁸ Similarly, marked differences in template quantity (the overall size of the sample) or the level of degradation¹⁹⁹ of the sample could generate doubts about the validity of the extrapolation. Allowing the opponent access to an automated forensic technique's source code would be one of the most effective ways to enable the opponent to conduct an exploration of such questions.

2. *Countervailing Considerations Against Granting the Defense Access*

Although the defense may have a significant interest in examining the source code of the forensic technique's computer software, the software's developer is likely to counter that it has an equally legitimate, countervailing interest in protecting the source code as a trade secret. Courts in the first wave of source code cases involving breath

196. See *People v. Robinson*, 860 N.Y.S.2d 159, 166 (App. Div. 2008).

197. One validation study for another probabilistic genotyping software program, STRmix, notes that a protocol had been validated only for one to three contributors. See STRmix V2.0.6 BFS CASEWORK INTERNAL VALIDATION SUMMARIES 2 (2016), <https://epic.org/state-policy/foia/dna-software/EPIC-16-02-02-CalDOJ-FOIA-20160219-STRmix-V2.0.6-Validation-Summaries.pdf>.

198. 2 GIANNELLI ET AL., *supra* note 5, § 18.04[b][2].

199. For that matter, various parts of a mixture may be degraded to a different degree. The various parts might involve different tissue types, or they have been deposited at different times. NICHE VISION SOLE SOURCE JUSTIFICATION: STRMIX MIXTURE INTERPRETATION SOFTWARE—JUSTIFICATION FOR NON-COMPETITIVE PROCUREMENT 2–3 (2014), <https://epic.org/state-policy/foia/dna-software/>. In the REPORT TO THE PRESIDENT—FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (2016), the President's Council of Advisors on Science and Technology states:

[C]urrent studies have adequately explored only a limited range of mixture types (with respect to number of contributors, ratio of minor contributors, and total amount of DNA). The two most widely used methods (STRmix and TrueAllele) appear to be reliable within a certain range, based on the available evidence and the inherent difficulty of the problem. Specifically, these methods appear to be reliable for three-person mixtures in which the minor contributor constitutes at least 20 percent of the intact DNA in the mixture and in which the DNA amount exceeds the minimum level required for the method.

Id. at 80.

testing devices often noted that the manufacturer involved in the case had asserted the privilege protecting its trade secret.²⁰⁰ The courts reasoned that defendants could not obtain discovery from the government because the government did not own or possess the code and that the defendant could not have discovery from the manufacturer due to the privilege.²⁰¹ In the second wave of case law involving TrueAllele, Cybergenetics similarly persuaded the courts that the privilege for its trade secrets trumps the defendant's discovery interests.²⁰²

Admittedly, a manufacturer that has invested time and money in the development of trade secret information has a legitimate interest in safeguarding its investment.²⁰³ If the information is publicly circulated and copied, the company can lose licensing revenue.²⁰⁴ If the software in question is one of the company's most valuable assets, the result might be the bankruptcy of the company. It is estimated that the theft of intellectual property costs American businesses \$300 billion annually.²⁰⁵ The protection of trade secrets is such a vital national interest that in 1996, Congress enacted the Economic Espionage Act,²⁰⁶ criminalizing the theft of trade secrets.²⁰⁷ Since the problem persisted, Congress passed the Theft of Trade Secrets Clarification Act of 2012.²⁰⁸ The Advisory Committee that proposed the Federal Rules of Evidence included draft Rule 508 which would have codified a privilege for trade secrets.²⁰⁹ Although Congress balked at enacting the draft rule, many states have done so; regardless, the federal courts have recognized the privilege by common-law process under Federal Rule 501.²¹⁰

200. See, e.g., *Moe v. State*, 944 So. 2d 1096, 1097 (Fla. Dist. Ct. App. 2006); *People v. Robinson*, 860 N.Y.S.2d 159, 167 (App. Div. 2008); *State v. Burnell*, 2007 WL 241230, at *2 (Conn. Super. Ct. Jan. 18, 2007); *People v. Cialino*, 831 N.Y.S.2d 680, 681 (Crim. Ct. 2007).

201. See, e.g., *Moe*, 944 So. 2d at 1097; *Robinson*, 860 N.Y.S.2d at 161; *Burnell*, 2007 WL 241230, at *2; *Cialino*, 831 N.Y.S.2d at 681.

202. David Kravets, *Secret Source Code Pronounces You Guilty as Charged*, ARS TECHNICA (Oct. 17, 2015, 7:00 AM), <https://arstechnica.com/tech-policy/2015/10/secret-source-code-pronounces-you-guilty-as-charged>; Palazzolo, *supra* note 27. In *Commonwealth v. Robinson*, the court tersely commented that “[t]his source code is the intellectual property of Cybergenetics.” No. CC 201307777, at *1 (Pa. Commw. Ct. Allegheny Cty. Feb. 4, 2016).

203. Mislow, *supra* note 12, at 40–41 (discussing research and development costs).

204. *Id.* at 41.

205. Rachel Glas, *Trade-Secret Cases Abound in California*, CAL. LAW., May 2011, at 6.

206. Pub. L. 104-294, 110 Stat. 3488 (codified as amended at 18 U.S.C. §§ 1831–1839 (2012)).

207. 18 U.S.C. § 1831; Joseph F. Savage, Jr. et al., *Trade Secrets: Conflicting Views of the Economic Espionage Act*, CRIM. JUST. MAG., Fall 2000, at 11.

208. Theft of Trade Secrets Clarification Act of 2012, Pub. L. No. 112-236, 126 Stat. 1627 (codified as amended at 18 U.S.C. §§ 1831–1839 (2012)).

209. IMWINKELRIED, *THE NEW WIGMORE*, *supra* note 101, at 1456.

210. *Id.*

Although companies like CMI and Cybergenetics have a perfect right to assert the evidentiary privilege for their trade secrets, their assertion of the right should not end the analysis. Under statute and at common law, it is well-settled that the privilege is conditional or qualified one.²¹¹ Draft Federal Rule of Evidence 508 expressly described the privilege as “qualified.”²¹² A party seeking discovery of the trade secret information can defeat the privilege claim by demonstrating that the information is highly relevant and necessary for trial.²¹³

While a judge might assign considerable weight to the company’s interest in safeguarding its trade secrets, she can take steps to reduce the risks posed by the disclosure of the information during pretrial discovery. The second sentence of draft Rule 508 read: “When disclosure is directed, the judge shall take such protective measures as the interests of the holder of the privilege and of the parties and the furtherance of justice may require.”²¹⁴ In civil cases, courts have issued the following protective orders, *inter alia*: The opposing party’s experts could examine the trade secret information only in a secure room;²¹⁵ to gain access to the secure room, the experts had to identify themselves by iris and palm-print scans;²¹⁶ during their examination of the information, the experts had to use paper bearing tags emitting radio waves to determine how many pages of notes the experts had used;²¹⁷ counsel and the experts had to sign declarations that they would access the data only for use in the present litigation;²¹⁸ and the trial courtroom would be closed to the public during any testimony discussing the trade secret information.²¹⁹ Some commentators generalize that when the trade secret information is important enough in the case, the civil courts almost always issue protective orders and re-

211. *Farouk Systems, Inc. v. Costco Wholesale Corp.*, 700 F. Supp. 2d 780, 788 (S.D. Tex. 2010); *Davis v. Leal*, 43 F. Supp. 2d 1102, 1110 (E.D. Cal. 1999).

212. *IMWINKELRIED, THE NEW WIGMORE*, *supra* note 101, at 1471 (quoting *FED. R. EVID.* 508).

213. *Bridgestone Ams. Holding Inc. v. Mayberry*, 878 N.E.2d 189, 195 (Ind. 2007). There is necessity if the party seeking discovery cannot obtain equivalent information through reasonably available means.

214. *IMWINKELRIED, THE NEW WIGMORE*, *supra* note 101, at 1471.

215. Amanda Bronstad, *Security Tight as Toyota Shares Its “Crown Jewels,”* NAT’L L.J. (Mar. 7, 2011), <http://www.nationallawjournal.com/=1202484326281?keywords=Toyota+Crown+Jewels&publication=National+Law+Journal>.

216. *Id.*

217. *Id.*

218. *Bernier v. One World Techs., Inc.*, 746 F. Supp. 2d 240, 244–45 (D. Mass. 2010).

219. Peter Page, *Tire Trial Tactic: Clear the Courtroom*, NAT’L L.J., Oct. 29, 2002, at A4; *see also Foster & Northrop*, *supra* note 35, at 45 (describing protective orders for keeping source code confidential during discovery and trial).

quire disclosure.²²⁰ One court has observed that in civil practice, the real rule is that when a party seeking discovery has a genuine need for trade secret information, the technically privileged information is discoverable under protective order.²²¹

The courts have also required disclosure subject to protective order in criminal cases.²²² As previously stated, the theft of trade secrets is now a federal crime.²²³ In a prosecution, the defense might contend that the information in question does not qualify as a trade secret under federal law. To press that contention, a defendant may seek discovery of the content of the alleged trade secret. As in civil cases, criminal courts have balanced the competing interests—the accused’s need for information and the trade owner’s legitimate interest in safeguarding the information—by issuing protective orders. The courts have entered pretrial pre-indictment protective orders, pretrial post-indictment orders, and trial orders.²²⁴ As in civil cases, courts have crafted myriad of orders, including pretrial provisions that: the experts to be granted access undergo certain vetting; the experts sign a declaration acknowledging their obligation not to circulate the information; the expert be allowed to study the information only in secure areas; the experts be obliged to conduct their analysis on protected computers; and the experts be required to return or destroy all notes as soon as the investigation or prosecution terminates.²²⁵ At trial, judges have given jurors special instructions about their obligation not to disclose trade secrets discussed during testimony in open court and entered orders dealing with such matters as the redaction of trial exhibits, the filing of exhibits under seal, and the closure of the courtroom during testimony mentioning the trade secrets.²²⁶

Given these precedents, the assertion of a trade secret privilege by companies, such as CMI and Cybergenetics, should not bar pretrial discovery. In published civil opinions, courts have sometimes ordered disclosure to counsel representing, or experts hired by, litigants with economic interests directly adverse to those of the owner of the trade

220. 26 CHARLES A. WRIGHT & KENNETH W. GRAHAM, JR., *FEDERAL PRACTICE AND PROCEDURE* § 5650 (2d ed. 1992) (“[O]nce necessity is shown, the courts will almost always order disclosure.”).

221. *La Chemise Lacoste v. General Mills, Inc.*, 53 F.R.D. 596, 605 (D. Del. 1971), *aff’d*, 487 F.2d 312 (3d Cir. 1973).

222. See Brian L. Levine & Timothy C. Flowers, *Your Secrets Are Safe with Us: How Prosecutors Protect Trade Secrets During Investigation and Prosecution*, 38 AM. J. TRIAL ADVOC. 461 (2015).

223. 18 U.S.C. §§ 1831–1839 (2012) (criminalizing theft of trade secrets).

224. Levine & Flowers, *supra* note 222, at 480–81.

225. *Id.* at 481.

226. *Id.* at 481–82.

secret.²²⁷ In that setting, the risk of wrongful disclosure is at its highest. The court is ordering divulgence to a litigant or expert hired by a litigant who is the owner's actual or potential economic competitor—someone with the greatest temptation to misuse the information to the owner's detriment.²²⁸ In sharp contrast, in the typical case in which the court orders disclosure to an accused challenging an automated forensic science technique, the risk is much smaller. The accused is not an economic competitor of the owner. The defense's only interest is using the trade secret information in the pending litigation to raise doubts about the reliability of the computerized technique. If protective orders suffice to protect the owner's interests in cases ordering disclosure to the owner's business rivals, there is all the more reason to mandate disclosure here.

3. *A Proposal for Balancing the Defense's Need for Access Against the Countervailing Considerations*

Some might suggest that the court should automatically order discovery of the source code whenever an expert relies on an automated technique. However, this Article advances a much less sweeping claim. Faced with competing legitimate interests, a trial judge must attempt to strike a rational balance. In this context, the judge could do so by proceeding in two steps. First, a judge should assign to the accused seeking discovery the burden of showing that the facts of the instant prosecution exceed, or are at the margins of, the validation range of the empirical studies relied on by the prosecution. More specifically, the defendant must convince the judge that the available studies do not adequately address the effect of a specified, material variable or condition present in the instant case. The most clear-cut case would be a fact situation in which none of the available studies relied on by the prosecution experts tested the application of the technique to fact situations involving the condition.

The judge should certainly not accept the *ipse dixit* assertion of the defense counsel that the omitted condition is material in the sense that its presence could affect the outcome of the test. After all, the defense counsel is not a scientist. Rather, the judge ought to demand that the defense present expert testimony explaining why it is plausible that that condition could change the test result. In modern DNA typing, analysts evaluate the peaks on electropherograms to identify the alleles present in a sample.²²⁹ The height of the peaks is measured

227. *Id.* at 481.

228. *Id.* at 480–81.

229. 2 GIANNELLI ET AL., *supra* note 5, § 18.04[b][2].

in relative fluorescent units.²³⁰ In mixed samples, the peak reflecting one contributor's DNA marker at a site might be thirty units in height while those reflecting another contributor's profile could be twice as tall.²³¹ Suppose that in the case of computer software such as TrueAllele, the prosecution's expert relied on validation studies in which there was at most a 3:1 ratio between the height of the peak attributable to the largest contributor and that attributable to the smallest contributor. In the instant case, the disparity ratio was 5:1. The judge should not accept at face value the defense counsel's claim that the difference between a 3:1 ratio in the studies and a 5:1 ratio in the pending case invalidates the extrapolation from the studies to the opinion that the technique can be reliably applied in the pending case. Instead, the judge ought to insist that the defense produce an expert²³² explaining why that difference could plausibly impact the outcome of the analysis. Unless the defense could do so, the judge ought to deny the defense discovery request.

Assume that in the first step, the judge concludes that the defense has met its burden. Even then the judge should not automatically require the manufacturer to furnish the defense with a printout or electronic version of the source code. Instead, the judge could give the manufacturer a choice to: either (1) allow the defense to test the application of the program to a fact situation including the material condition or variable omitted from the validation studies, or (2) provide the defense with the source code. At the first step in the analysis, the defense criticized the empirical record on the ground that the validation studies conducted to date omit a material condition. Enabling a defense expert to design and conduct a new validation study covering that condition directly responds to such criticism. Significantly, Cybergenetics "has offered to let defense lawyers conduct their own tests on the software to satisfy themselves of its reliability."²³³ At the end of this first step, the judge is not licensing a fishing expedition of unlimited scope; rather, the judge is authorizing discovery designed to meet a discrete defense criticism of the state of the empirical record in order to determine whether the technique can be reliably applied to the facts in the pending case. By enabling the defense expert to conduct a new validation study testing that application,²³⁴ the manufac-

230. *Id.* § 18.04[b][1].

231. *Id.*

232. Of course, an indigent defendant may need to seek funds from the court to hire an expert and cover the costs of testing. *Id.* §§ 4.03–4.05.

233. Palazzolo, *supra* note 27.

234. Again, an indigent defendant might need to seek government funds to hire a defense expert. 2 GIANNELLI ET AL., *supra* note 5, §§ 4.03–4.05.

turer would afford the defense expert a fair opportunity to investigate the merit of the criticism.

If, for some reason, the manufacturer rejected the first option, the judge could then order the manufacturer to grant the defense access to the source code itself.²³⁵ Since the code usually embodies the manufacturer's trade secret, presumably the manufacturer would choose the first option. Either option, though, would suffice to enable the defense to reasonably determine whether the inclusion of the additional condition could actually—not merely theoretically—affect the outcome of the use of the automated forensic technique.

IV. CONCLUSION

Students of the criminal justice system have increasingly exposed the dangers of reliance on eyewitness testimony²³⁶ and confession evidence.²³⁷ To reduce its reliance on those suspect types of evidence, the criminal justice system has turned to forensic science techniques. Admittedly, there have been serious criticisms of the extent of the validation of several of those techniques;²³⁸ however, forensic techniques resting on solid validation promise to enhance the objectivity and accuracy of the evidence on which criminal verdicts are based.

The justice system's increasing reliance on forensic evidence has placed growing demands on forensic laboratories.²³⁹ DNA testing is a case in point. In many jurisdictions, there are large backlogs of samples, including rape test kit samples, that have yet to be processed.²⁴⁰ The problem has reached such acute proportions that in 2000 Con-

235. The defense expert's review could be a costly, time-consuming process. See Foster & Northrop, *supra* note 35, at 42, 44–45 (explaining that there could be “millions of lines of code along with hundreds—if not thousands—of interconnected files,” and a thorough review of the code could “take months”).

236. 2 GIANELLI ET AL., *supra* note 5, § 9.02 (discussing the weaknesses in lay eyewitness testimony).

237. *Id.* § 9.09 (discussing false confessions and the weaknesses in confession evidence).

238. See generally NAT'L RES. COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 100 (2009) (explaining DNA analysis is the only “forensic method” that “has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between an evidentiary sample and a specific individual or source”).

239. Dawson, *supra* note 11, at 10.

240. ANDRE A. MOENSSENS ET AL., SCIENTIFIC EVIDENCE IN CIVIL AND CRIMINAL CASES § 17.14 (5th ed. 2007) (“[O]ne major hurdle left to be resolved is the backlog of cases waiting to be analyzed” and “[w]ith the increased submissions of DNA samples,” the system can be “overwhelmed”).

gress felt compelled to pass the DNA Backlog Elimination Act²⁴¹ to provide additional funding to reduce the backlog.²⁴²

In both the public and private sectors, the response to backlog is often technological automation.²⁴³ One of the major breakthroughs in the history of fingerprint examination was the advent of automated fingerprint identification systems which dramatically shortened the time required for a fingerprint analysis by eliminating the need for time-consuming manual sorting and classification of prints.²⁴⁴ Of course, eventually DNA typing replaced fingerprinting as the “gold standard” of forensic science. Initially DNA analyses were conducted manually, but today most laboratories in the United States utilize automated Applied Biosystems instrumentation or similar equipment.²⁴⁵ When it became evident that one of the greatest challenges in DNA analysis was the evaluation of mixed samples, Cybergenetics developed another level of automation, namely, TrueAllele software.²⁴⁶ As we have seen, the key to automation is the development of source code.²⁴⁷ There cannot be automation without some form of source code.²⁴⁸

The hope is that increasing automation of forensic science techniques, such as infrared breath testing and the assessment of mixed DNA samples, will not only increase the efficiency of forensic laboratories but also enhance the accuracy of the forensic analyses introduced in court. In its 2016 report on forensic science, the President’s Council of Advisors on Science and Technology asserts that the advent of “probabilistic genotyping software programs clearly represent[s] a major improvement over” the subjective methods of analyzing complex mixtures that have been employed in the past.²⁴⁹ That hope can be realized only if the source code for automated techniques is reliable. There are ways of indirectly gauging the reliability

241. Pub. L. No. 106-546, 114 Stat. 2726 (codified as amended at 42 U.S.C. § 14135 (2012)). In enacting the DNA Backlog Elimination Act, Congress made an explicit legislative finding that there is “a nationwide backlog of DNA samples from . . . crime scenes that need to be tested.” *Id.* § 14135.

242. *Id.* § 14135.

243. Harman, *supra* note 1.

244. MOENSSENS ET AL., *supra* note 240, § 10.06.

245. William Thompson & Daniel Krane, *DNA in the Courtroom*, in *PSYCHOLOGICAL AND SCIENTIFIC EVIDENCE IN CRIMINAL TRIALS* § 11:23 (2003) (noting 85% of American DNA laboratories employ such automation).

246. See Moss, *supra* note 23, at 1059.

247. Harman, *supra* note 1, at 3.

248. *Id.*

249. PRES. COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT – FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS 79 (2016), <http://pccinc.org/wp-content/uploads/2016/10/Report-to-President.pdf>.

of the source code without directly examining the code itself. Empirical studies can provide a circumstantial assurance of the analyses' trustworthiness so long as the application of the technique in the case falls within the validation range established by studies. When the conditions of the case are beyond, or at the edges of, the validation range, courts should grant the defense pretrial discovery to investigate whether the omitted condition could change the outcome of the automated test. The discovery can take the form of either the disclosure of the source code, or according to defendants an opportunity to conduct a new validation test covering the omitted condition. Until courts guarantee the defense that right, source code will continue to be a source of controversy and doubt about the marked trend toward the automation of forensic analysis in the United States.